

University of Nebraska - Lincoln

**DigitalCommons@University of Nebraska - Lincoln**

---

US Department of Energy Publications

U.S. Department of Energy

---

2011

## Methods for probabilistic modeling of concentrating solar power plants

Clifford Ho

*Concentrating Solar Technologies, Albuquerque, NM*

Siri S. Khalsa

*Concentrating Solar Technologies, Albuquerque, NM*

Gregory J. Kolb

*Photovoltaics and Grid Integration, Albuquerque, NM*

Follow this and additional works at: <https://digitalcommons.unl.edu/usdoepub>



Part of the [Bioresource and Agricultural Engineering Commons](#)

---

Ho, Clifford; Khalsa, Siri S.; and Kolb, Gregory J., "Methods for probabilistic modeling of concentrating solar power plants" (2011). *US Department of Energy Publications*. 115.

<https://digitalcommons.unl.edu/usdoepub/115>

This Article is brought to you for free and open access by the U.S. Department of Energy at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in US Department of Energy Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.

# Methods for probabilistic modeling of concentrating solar power plants

Clifford K. Ho<sup>a,\*</sup>, Siri S. Khalsa<sup>b</sup>, Gregory J. Kolb<sup>c</sup>

<sup>a</sup> Sandia National Laboratories, Concentrating Solar Technologies, P.O. Box 5800, Albuquerque, NM 87185-1127, USA

<sup>b</sup> Sandia Staffing Alliance, Concentrating Solar Technologies, P.O. Box 5800, Albuquerque, NM 87185-1127, USA

<sup>c</sup> Sandia National Laboratories, Photovoltaics and Grid Integration, P.O. Box 5800, Albuquerque, NM 87185-1127, USA

Received 30 November 2009; received in revised form 26 March 2010; accepted 6 May 2010

Available online 2 June 2010

Communicated by: Associate Editor Juergen Dersch

## Abstract

Probabilistic modeling of concentrating solar power technologies provides important information regarding uncertainties and sensitivities not available from deterministic models. Benefits of using probabilistic models include quantification of uncertainties inherent in the system and characterization of their impact on system performance and economics. This paper presents the tools necessary to conduct probabilistic modeling of concentrating solar technologies. The probabilistic method begins with the identification of uncertain variables and the assignment of appropriate distributions for those variables. Those parameters are then sampled using a stratified method (Latin Hypercube Sampling) to ensure complete and representative sampling from each distribution. Models of performance, reliability, and/or cost are then simulated multiple times using the sampled set of parameters. The results yield a cumulative distribution function that can be analyzed to quantify the probability of achieving a particular metric (e.g., net energy output or levelized energy cost) and to rank the importance of the uncertain input parameters.

© 2010 Elsevier Ltd. All rights reserved.

**Keywords:** Probabilistic modeling; Uncertainty; Sensitivity; Latin Hypercube Sampling; Concentrating solar

## 1. Introduction

Modeling system performance and economics of solar thermal power plants has traditionally relied on deterministic analyses. Input parameters are typically entered as specific (point) values rather than distributions of values that honor the inherent uncertainty in many of the system features and processes. As a result, the confidence and uncertainty associated with the results are unknown.

This paper introduces probabilistic tools to yield uncertainty analyses that can quantify the impact of system uncertainties on the simulated performance metrics. The confidence and likelihood of the simulated metric (e.g., annual energy produced, levelized energy cost) being above or below a particular value or within a given range

can be readily assessed and presented using these probabilistic methods. In addition, sensitivity analyses can be used with probabilistic analyses to determine the most important components, features, and/or processes that impact the simulated performance. This information can be used to guide and prioritize future research and characterization activities that are truly important to the relevant performance metrics.

## 2. Modeling approach

The probabilistic modeling approach consists of three primary steps: (1) creating uncertainty distributions for stochastic parameters and sampling the distributions  $n$  times, (2) running the performance and/or cost models  $n$  times using the sampled variables, and (3) evaluating the distribution of  $n$  results to quantify uncertainty and sensitivity. Screening analyses are first conducted to determine a

\* Corresponding author. Tel.: +1 505 844 2384.

E-mail address: [ckho@sandia.gov](mailto:ckho@sandia.gov) (C.K. Ho).

subset of input parameters that are to be assigned uncertainty distributions as opposed to deterministic point values. The uncertainty distributions (e.g., uniform, normal) can be based on actual data, literature values, or professional judgment. Monte Carlo or Latin Hypercube Sampling methods are then implemented in the model to generate many different (but equally probable) realizations of the system performance. Latin Hypercube Sampling requires fewer realizations than Monte Carlo sampling, which is prone to clustering, and allows prescribed correlations among parameters. The ensemble of realizations generates a cumulative probability distribution that can be used to quantify the uncertainty in system performance. A stepwise regression analysis is then performed to determine the input parameters that are most correlated to the simulated performance metric, indicating those parameters or processes that are most important to the system performance. These types of analyses provide additional useful information not available in deterministic analyses (or even parametric analyses where only a few prescribed values are varied in a “one-off” fashion). Additional details of each of the three primary steps is provided in the next few sections.

### 2.1. Uncertainty distributions and sampling

Many parameters used in models of concentration solar power technologies are not known precisely because of a lack of measurement data, natural variability in the parameter value (e.g., mirror reflectivity), or changes in future behavior (e.g., insolation). To accommodate this inherent uncertainty, parameter distributions can be created to represent a range of values for each uncertain parameter. Latin

Hypercube Sampling (LHS) is a way to sample these distributions in a systematic (stratified) way to ensure that values are sampled from across the entire distribution with the chosen number of realizations,  $n$ . Fig. 1 shows examples of how parameters with normal and uniform distributions would be stratified for  $n = 5$  samples. The cumulative distribution function (CDF) shown on the right side of Fig. 1 is stratified into five equal bins along the  $y$ -axis, where the letters along the  $x$ -axis represent the values of the parameter. Because a normal distribution has a higher density of values near the mean, the bins are not distributed equally along the  $x$ -axis (i.e., a smaller range of values covers a greater range of probability near the mean). In contrast, a uniform distribution yields equally spaced bins throughout the parameter distribution.

A code has been written by Sandia National Laboratories to implement LHS (Wyss and Jorgensen, 1998). After sampling the uncertain variables  $n$  times, the code can also produce restricted pairings such that each set of parameters honors correlations (or zero correlations) among the sample variables. For example, the magnitude of the incident flux on the receiver may be inversely correlated to wind speed (wind shakes the collectors and lowers the optical intercept). These correlations can be specified in LHS, and the code will swap the order of the  $n$  parameters until the desired correlations exist among the sampled parameters. The minimum number of samples required to implement a restricted pairing among the sampled variables (either to correlate variables or to minimize correlation) is approximately  $4k/3$ , where  $k$  is the total number of uncertain variables (Wyss and Jorgensen, 1998).

Typically, the number of samples (realizations) that are used to represent all possible input-parameter combina-

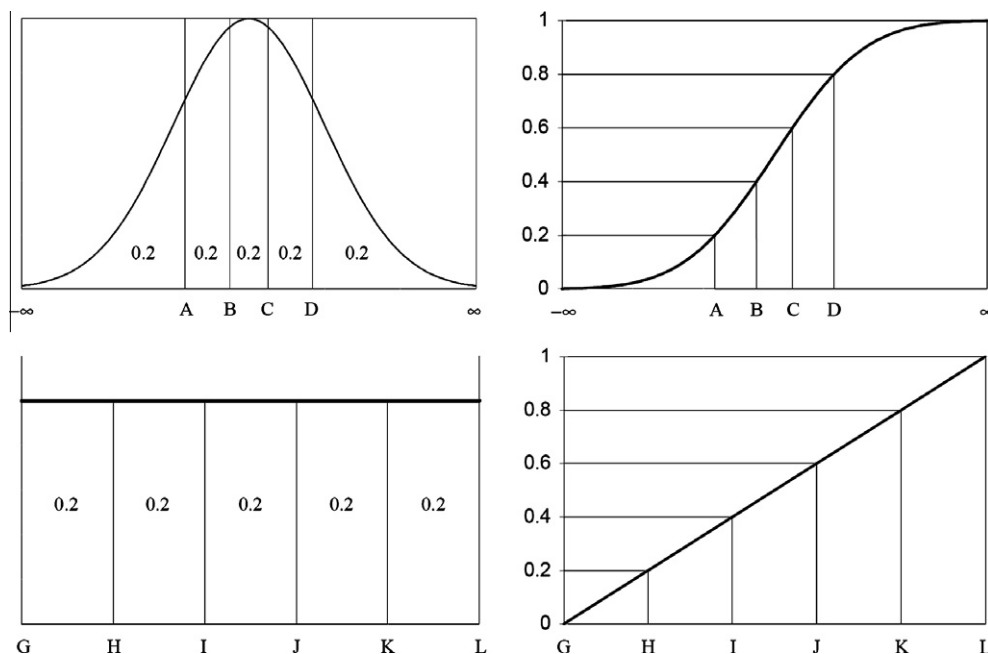


Fig. 1. Histograms (left) and CDFs (right) of parameters with a normal distribution (top) and uniform distribution (bottom) stratified into five equally probable bins (from Wyss and Jorgensen, 1998).

tions in the model is increased until the cumulative distribution function of the simulated metric no longer changes. This indicates that the variability in the simulated metric caused by uncertainty in the input parameters is adequately covered by the sampled parameter combinations. Alternatively, Section 2.3 provides a method of estimating the number realizations that are needed to provide a quantified confidence that the simulated mean is within a prescribed tolerance of the true mean for the simulated metric. However, the tails of the cumulative distribution function are likely to be more sensitive to changes in the sample size than the mean. So, increasing the number of samples used in the model until the entire cumulative distribution function (especially the tails) no longer changes is recommended.

## 2.2. Probabilistic modeling using SOLERGY and other performance or cost models

For each set of sampled input parameters, a model is applied to simulate the desired cost or performance metric. For example, if 100 values are sampled from each uncertain input-parameter distribution, the model would be run 100 times to yield 100 simulated results. The paragraphs below describe methods for applying models and codes in this probabilistic (multi-realization) mode.

SOLERGY (Stoddard et al., 1987) simulates the annual energy output of a solar thermal central receiver power plant and has been validated using data from Solar One (Alpert and Kolb, 1988). It utilizes actual or simulated weather data at time intervals of 15 min and calculates the net electrical energy output at every time step throughout an entire year. Input to the code is entered via user-specified text files.

Factors include energy losses in each component of the system, delays incurred during start-up, weather conditions, storage strategies, and power limitations for each component. A solar power plant simulated by SOLERGY is defined by the set of parameters contained in a SOLERGY input file. The weather data used in a SOLERGY simulation is listed in a weather file, which contains direct normal insolation, wind direction, wind speed, dew point temperature, barometric pressure, and dry bulb temperature listed every 15 min for 1 year.

LHS produces a matrix of parameter values, where each row corresponds to one realization, and each column contains a parameter representing one factor of the performance of the solar power plant. SOLERGY Batch Mode is a program that has been recently developed to run SOLERGY in a probabilistic mode. SOLERGY Batch Mode runs all the realizations defined in an LHS output matrix and automatically compiles desired output values, e.g. Net Energy Output, from all realizations. A flowchart illustrating the operation of SOLERGY Batch Mode is displayed in Fig. 2.

The SOLERGY Batch Mode Input File contains the list of parameter values for each realization (i.e. the LHS output matrix). The Template SOLERGY Input File is a modified SOLERGY Input File, with placeholders substituted for the locations at which parameter values must be inserted. SOLERGY Batch Mode reads the parameter values for the first realization, and then substitutes these values for the placeholders in the Template SOLERGY Input File, thus producing a SOLERGY Input File corresponding to this realization. The entire matrix of realizations is processed in this manner. After all SOLERGY Input Files are created, SOLERGY Batch Mode sequentially passes each input file

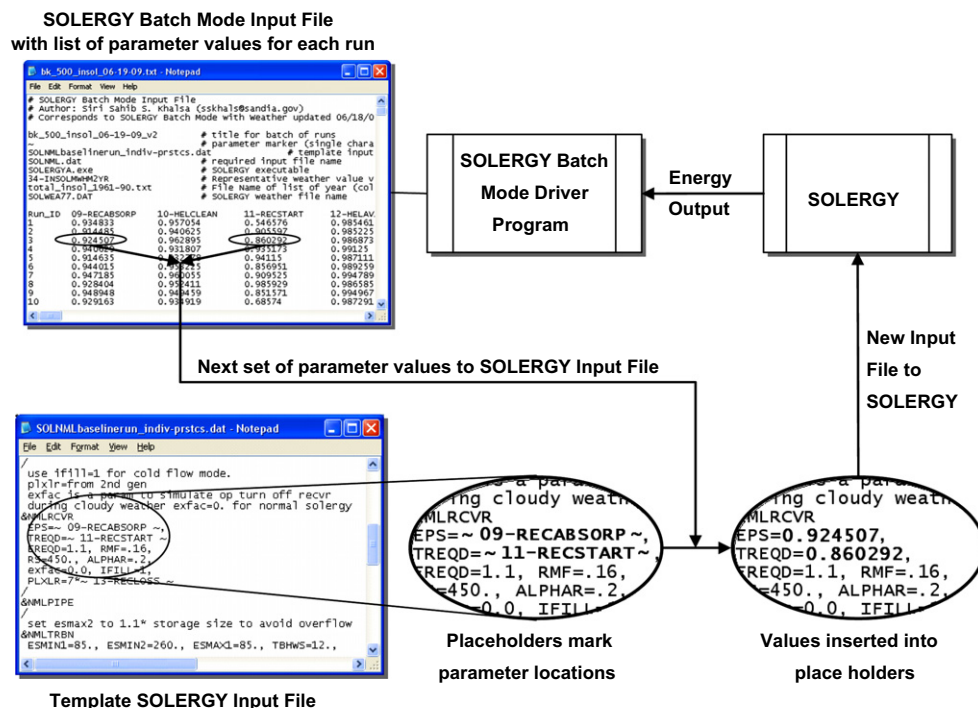


Fig. 2. Flowchart illustrating the operation of SOLERGY Batch Mode.

to SOLERGY for simulation. Finally, SOLERGY Batch Mode compiles the Net Energy Output corresponding to each realization.

SOLERGY Batch Mode can also simulate weather uncertainty by accepting a list of values from LHS for a parameter that characterizes an entire year of weather data (e.g., total annual direct normal insolation), select the year of weather data that is best characterized by this parameter value, and then pass the selected weather file to SOLERGY to be used with the corresponding realization's SOLERGY Input File. An example utilization of this capability of simulating weather uncertainty is discussed in Section 3.1.

Any model can be run in probabilistic mode using the method of passing multiple sets of input parameter values generated from an LHS output matrix to the model, as described above with the example of SOLERGY. Models with a command-line interface could be made to run in probabilistic mode in the same manner illustrated in Fig. 2.

Models that use a Graphical User Interface (GUI) for input without an available command-line interface, such as the Solar Advisor Model (SAM) by NREL (Gilman et al., 2008), can be run in probabilistic mode by gaining access to the model's input functions via a Dynamic Link Library (DLL). The parameter values listed in the LHS output matrix would be passed to the model by a driver program similar to SOLERGY Batch Mode, but through the DLL functions instead of through an input file passed to the model on the command-line. A DLL is created by compiling the model's source code with the option of making certain functions available for use through the DLL. Efforts are underway to implement LHS and probabilistic methods in SAM.

### 2.3. Uncertainty and sensitivity analyses

After the  $n$  realizations are simulated, the resulting distribution of results can be analyzed by using a cumulative distribution function to quantify the likelihood of achieving a cost or performance metric. For example, one may wish to use the resulting CDF to quantify the uncertainty associated with the simulated output parameter (e.g. net annual energy output, levelized cost of electricity), or to quantify the probability of achieving a prescribed performance or cost metric.

Prior to utilizing the CDF for uncertainty and probability assessments, one needs to assess the sufficiency of the number of realizations on the CDF output. A probabilistic model should generate a probability distribution that is representative of the population of possible outcomes that results from taking all uncertainties into consideration. Each realization of the model is one sample of this population. Therefore, the extent to which the simulated distribution of outputs represents the true population improves when the number of realizations increases. Before using a simulated distribution to calculate probabilities, it is important to quantify the confidence that the simulated distribution appropriately represents the true population of possible outcomes. One

approach is to determine the validity confidence,  $c$ , with which one can claim that the true population mean falls within, say,  $\pm 0.5\%$  of the simulated mean. This confidence,  $c$ , is then given by the probability:

$$c = P\left(-0.005\bar{x}\frac{\sqrt{n}}{s} < t_{n-1} < 0.005\bar{x}\frac{\sqrt{n}}{s}\right) \quad (1)$$

where  $n$  is the number of realizations,  $t_{n-1}$  is the variable of Student's  $t$ -distribution with  $n - 1$  degrees of freedom,  $\bar{x}$  is the simulated mean,  $s$  is the sample standard deviation of the simulated distribution of outputs, and a tolerance of  $0.5\%$  about the simulated mean is assumed. An example of the calculation of the validity confidence is provided in Section 3.2.

In addition to the uncertainty analysis described above, sensitivity analyses can be performed with probabilistic models to identify those input parameters that most impact the simulated performance metric. The sensitivity of the probabilistic model to uncertain input variables can be determined using regression analysis. Multiple regression analysis involves construction of a linear regression model of the simulated output (the dependent variable) and the stochastic input variables (independent variables) using a least-squares procedure. Stepwise linear (rank) regression is a modified version of multiple regression that selectively adds input parameters to the regression model in successive steps (Helton and Davis, 2000). In this method, a sequence of regression models is constructed that successively adds the most important input parameters to the regression to improve the overall correlation. In the end, the sensitivity analysis identifies those parameters that are significantly correlated to the performance metric, and omits those parameters that are not.

The sensitivity of the independent variables can be represented by  $\Delta R^2$ , the change in the coefficient of determination when a new independent variable is added to the model. The value of  $\Delta R^2$  describes the percentage of the uncertainty or variability in the simulated metric that is caused by the uncertainty in each input variable. In addition to  $\Delta R^2$ , the standardized regression coefficient ( $\beta$ ) is another statistical measure that evaluates the relative contributions of each input parameter to the magnitude of the dependent variable (as opposed to the variability of the dependent variable). The sign of  $\beta$  also gives the direction of correlation. The importance ranking of the independent variables using either  $\Delta R^2$  or  $\beta$  are typically the same.

### 3. Case study

A hypothetical 100 MW<sub>e</sub> molten-salt central receiver power-tower system with thermal storage (Fig. 3) was simulated using SOLERGY Batch Mode. The model assumed a storage capacity of 7 h and used weather data for Daggett, CA. The influence of uncertainty in the insolation values was investigated by sampling from 30 years of insolation data in the probabilistic analysis.



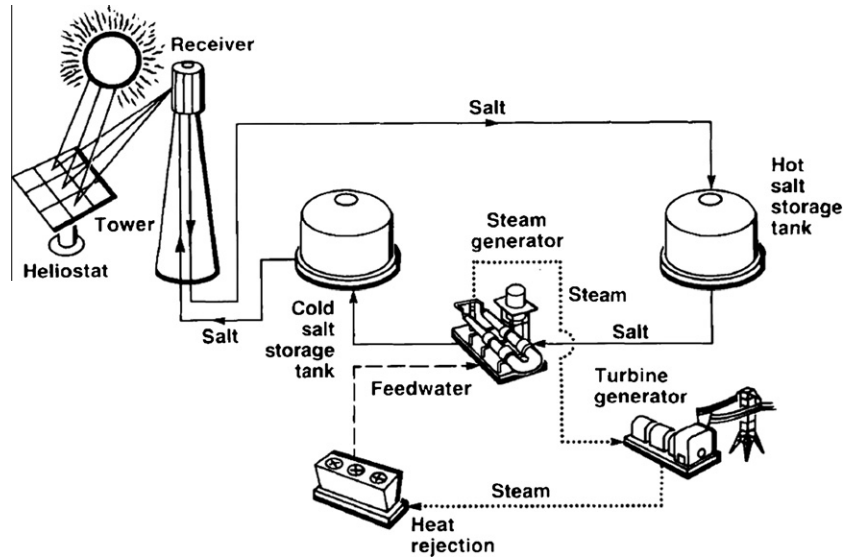


Fig. 3. Schematic of hypothetical molten-salt central receiver system with thermal storage.

### 3.1. Uncertainty distributions

The uncertainty distributions for the SOLERGY runs were taken from (Ho and Kolb, 2009) and are summarized in Table 1. The justifications for the parameter ranges are given in Ho and Kolb (2009).

Fig. 4 shows the CDF for the total annual insolation from 1961 to 1990 at Daggett, CA. This CDF was sampled using LHS, and the other parameters in Table 1 were sampled as well. Each total annual insolation value characterizes 1 year of meteorological data measured in Daggett, CA, (Weather Bureau Army Navy (WBAN) identification number 23161). The data was obtained from the National Solar Radiation Data Base (NSRDB) ([http://rredc.nrel.gov/solar/old\\_data/nsrdb/1961-1990/hourly/](http://rredc.nrel.gov/solar/old_data/nsrdb/1961-1990/hourly/)) in an hour-by-hour format for every year in the period January 1, 1961 to December 31, 1990. A program was written to extract the hourly direct normal insolation from the NSRDB hourly data files and create corresponding SOLERGY Weather Files. Each hour of NSRDB data was assumed to remain constant throughout the hour, as a SOLERGY Weather File lists data at 15 min intervals. The resultant SOLERGY Weather Files differed only by hourly direct normal insolation, and all contained the

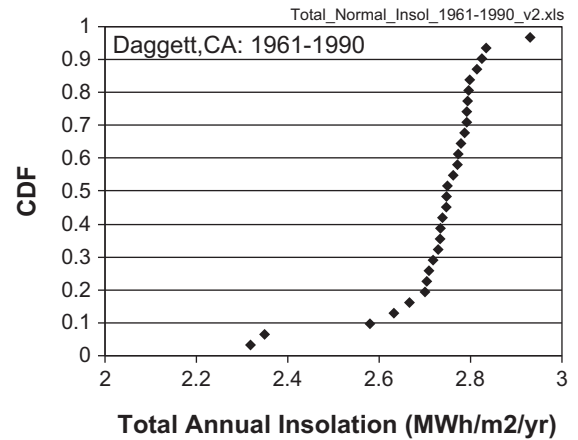


Fig. 4. CDF for the total annual insolation at Daggett, CA from 1961 to 1990.

same set of wind directions, wind speeds, dew point temperatures, barometric pressures, and dry bulb temperatures taken from 1977 data from Aerospace Corp.

The total annual insolation,  $I_{\text{total}}$ , in  $\text{MW h/m}^2/\text{yr}$  for a given year is defined as the total energy per unit area for the entire year received within a  $5.7^\circ$  field of view that is centered on the sun ([http://rredc.nrel.gov/solar/old\\_data/nsrdb/1961-1990/hourly/tab2.html](http://rredc.nrel.gov/solar/old_data/nsrdb/1961-1990/hourly/tab2.html)), and is given by

$$I_{\text{total}} = \sum_i DNI(t_i) \Delta t \quad (2)$$

where  $DNI(t_i)$  is the direct normal insolation in  $\text{W/m}^2$  at hour  $t_i$ ,  $\Delta t$  is equal to 1 h, and  $i$  ranges from the first to the last hour of the year.

### 3.2. Probabilistic modeling

Three cases with differing numbers of realizations (64, 200, and 500), but with parameter values generated from the same uncertainty distributions, were simulated using

Table 1

Uncertainty distributions for SOLERGY parameters (from Ho and Kolb, 2009).

Parameter	Central value	Distribution
Receiver absorptance	0.93	Uniform 0.91–0.95
Heliostat cleanliness	0.95	Uniform 0.93–0.97
Receiver start-up time	0.75 h	Uniform $\pm 33\%$
Heliostat availability	0.99	Uniform 0.985–0.995
Receiver thermal losses	26.2 MW	Uniform $\pm 24\%$
Parasitic multiplier	1	Uniform $\pm 20\%$
Total annual insolation	2.75 $\text{MW h/m}^2$	Discrete cumulative distribution (see Fig. 4)

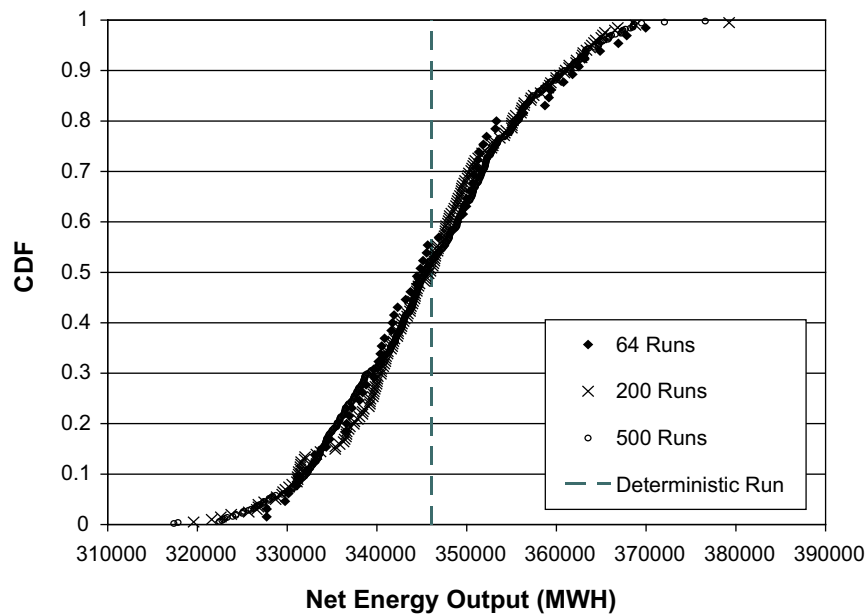


Fig. 5. Comparison of CDFs of net energy output for different numbers of realizations. Annual insolation held constant using 1977 data.

SOLERGY Batch Mode to obtain the annual net energy output (see Fig. 5). The total annual insolation was kept constant for this evaluation of different numbers of realizations. The respective validity confidences,  $c$ , described by Eq. (1) are presented in Table 2. The results show that the confidence that the true mean of all possible simulated net energy outputs falls within 0.5% of the simulated mean is 80.3%, 97.8%, and >99.9% for simulations with 64, 200, and 500 realizations, respectively. Therefore, 500 realizations was used to perform uncertainty and sensitivity analyses in the next section.

### 3.3. Uncertainty and sensitivity analyses

Once a sufficiently high confidence value has been obtained for the validity of a simulated CDF using the methods described in Section 3.2, the simulated CDF may be used to calculate probabilities for ranges of net energy output.

Fig. 6 displays the CDF for the Net Energy Output of the 500 realization probabilistic simulation described in Section 3.1. The 95% interval about the median, illustrated in Fig. 6, indicates that there is a 95% probability that the Net Energy Output will fall in the range 282,030–378,737 MW h/yr. Similarly, there is a 95% probability that the Net Energy Output will be at least 293,845 MW h/yr, and a 95% probability that the Net Energy Output will not exceed

371,019 MW h/yr. The deterministic result using central values from Table 1 is also shown in Fig. 6. The deterministic result is near the 50th percentile because all of the parameters (except insolation) were uniformly distributed.

Fig. 7 shows the results of a stepwise rank regression analysis of the 500 realizations. The standardized rank regression coefficient shows that the most important variable impacting the simulated net energy output was the annual insolation (positive correlation), followed by parasitics (negative correlation), receiver absorption (positive correlation), receiver heat loss (negative correlation), heliostat cleanliness (positive correlation), receiver start-up time (negative correlation), and heliostat availability (positive correlation). In Fig. 7, the plot of  $\Delta R^2$  shows that the variability in the annual insolation explains about 45% of the total variance in the simulated net energy output. Uncertainty in parasitics and receiver absorption both explain about 10% each of the total variance in simulated net energy output.

## 4. Summary

Probabilistic tools have been presented in this paper to assess the uncertainty and sensitivity of various parameters associated with concentrating solar power plants. Probabilistic modeling in this study consisted of three primary steps (1) sampling uncertainty distributions using LHS, (2) performing multiple simulations using a performance or cost model (e.g., SOLERGY), and (3) evaluating the uncertainty and sensitivity of the resulting distribution of simulated metrics. Although the analysis was focused on analysis of central receivers, the probabilistic methods and tools can be extended to other concentrating solar technologies as well. The results of these probabilistic analyses enables the analyst to quantify uncertainty associated with a simulated performance metric (e.g., net annual

Table 2  
Validity confidence values for simulations with different numbers of realizations.

Number of realizations, $n$	Standard deviation (MW h)	$t_{n-1}$	Validity confidence, $c$ (%)
64	10,616	1.303	80.28
200	10,619	2.303	97.77
500	10,909	3.544	99.96

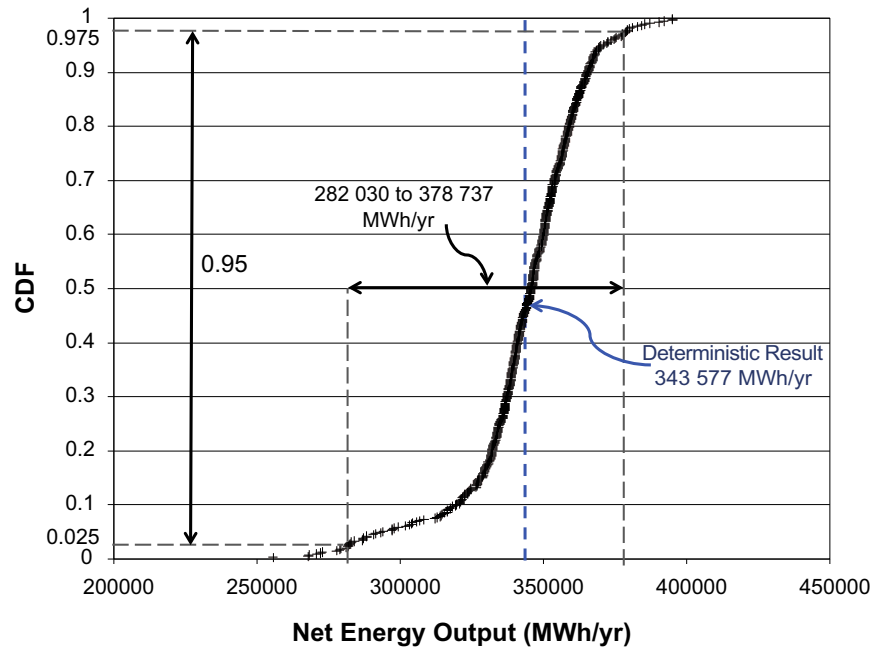


Fig. 6. CDF for simulated net energy output with 500 realizations. Annual insolation varied using data from 1961 to 1990.

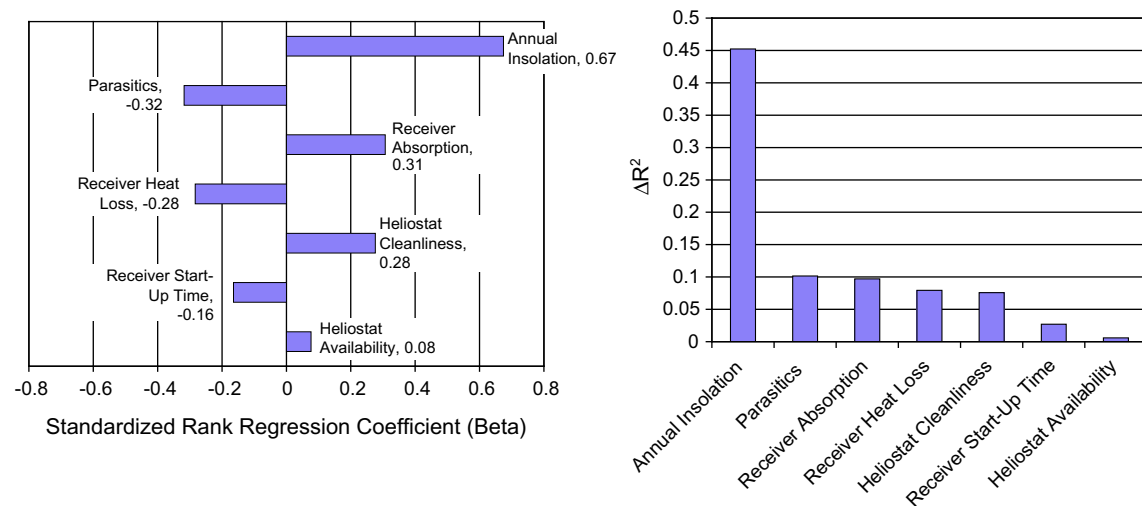


Fig. 7. Sensitivity analyses of uncertain parameters on simulated net energy output using standardized rank regression coefficient (left) and  $\Delta R^2$  (right).

energy output or levelized energy cost) and to identify which parameters are most important so that future research and characterization efforts can be prioritized.

### Acknowledgment

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

### References

Alpert, D.J., Kolb, G.J., 1988. Performance of the Solar One Power Plant as Simulated by the SOLERGY Computer Code. Sandia National Laboratories, Albuquerque, NM, SAND88-0321 (5 MB).

Gilman, P., Blair, N., Mehos, M., Christenson, C., Janzou, S., Cameron, C., 2008. Solar Advisor Model User Guide for Version 2.0. National Renewable Energy Laboratory Technical Report NREL/TP-670-43704, August 2008.

Helton, J.C., Davis, F.J., 2000. Sampling-Based Methods for Uncertainty and Sensitivity Analysis. Sandia National Laboratories, Albuquerque, NM, SAND99-2240 (5.3 MB).

Ho, C.K., Kolb, G.J., 2009. Incorporating uncertainty into probabilistic performance models of concentrating solar power plants. In: Proceedings of Energy Sustainability 2009. ES2009-90034, San Francisco, CA, July 19–23, 2009.

Stoddard, M.C., Faas, S.E., Chiang, C.J., Dirks, J.A., 1987. SOLERGY – A Computer Code for Calculating the Annual Energy from Central Receiver Power Plants. Sandia National Laboratories, Albuquerque, NM and Livermore, CA (SAND86-8060).

Wyss, G.D., Jorgensen, K.H., 1998. A User's Guide to LHS: Sandia's Latin Hypercube Sampling Software. Sandia National Laboratories, Albuquerque, NM, SAND98-0210 (560 KB).